

---

# From classroom tutor to hypertext adviser: an evaluation

Michael Kemp, H. C. Davis, William Roche and Wendy Hall  
The University of Southampton  
email: hcd@ecs.soton.ac.uk

---

*This paper describes a three-year experiment to investigate the possibility of making economies by replacing practical laboratory sessions with courseware while attempting to ensure that the quality of the student learning experience did not suffer. Pathology labs are a central component of the first-year medical undergraduate curriculum at Southampton. Activities in these labs had been carefully designed and they were supervised by lab demonstrators who were subject domain experts. The labs were successful in the eyes of both staff and students but were expensive to conduct, in terms of equipment and staffing. Year by year evaluation of the introduction of courseware revealed that there was no measurable difference in student performance as a result of introducing the courseware, but that students were unhappy about the loss of interaction with the demonstrators. The final outcome of this experiment was a courseware replacement for six labs which included a software online hypertext adviser. The contribution of this work is that it adds to the body of empirical evidence in support of the importance of maintaining dialogue with students when introducing courseware, and it presents an example of how this interaction might be achieved in software.*

---

## Foreword

The work in this paper was carried out by Michael Kemp as part of his Ph.D. studies under the supervision of William Roche in Pathology and Wendy Hall in the Department of Electronics and Computer Science, and with assistance from Hugh Davis and the Microcosm team.

The Ph.D. was examined in June 2000 by Hugh Davis (internal) and David Squires (external), shortly before the latter's diagnosis of cancer. At the time David had encouraged us to publish the work in *ALT-J*. We postponed the publication of this work so that it could appear in this memorial issue.

---

## Introduction

In response to an initiative to improve student attendance and appreciation of pathology practicals, case-based teaching was introduced to a first-year undergraduate Pathology course (McCullagh and Roche, 1992). The practicals were designed as part of a total experience, to build upon material recently presented in lectures, and they typically included notes on a case study, slides and a microscope. Students were required to answer a number of questions related to the case study and the practicals were followed by tutorial discussion groups in which the issues raised could be explored and reflected upon. The practicals were self-paced and the students could ask for assistance from demonstrators whenever they wished, although the majority of problems were sorted out by discussion between students.

These labs were informal; there were generally around eighty students present in each lab session, and there was a background noise consistent with many students engaging in conversation. Students arranged themselves informally into groups of between two and eight. These labs were perceived to be popular and effective. The problems were the difficulty in recruiting sufficient trained demonstrators and the potential for some students to take back seats. In some cases perhaps only one student would do the microscope work in spite of the possibility of using monitors.

However, there was a need to economize and it was decided to conduct a two-part experiment:

- to introduce some courseware to replace the practicals;
- to investigate the possibility of doing without demonstrators for the courseware practicals.

At the same time the pathology staff hoped that there might be educational gains from the introduction of courseware, since courseware may be used many times, allowing students the opportunity, after their tutorial discussion groups for example, to revisit the work and reconceptualize the material (Mayes, 1993; Mayes, 1995).

In this paper we start by briefly describing the research context within which this work was conducted. We continue by describing the nature of the courseware that was designed to replace the practicals and by detailing the results of the first evaluation of this courseware, which demonstrated that student learning did not suffer from the use of the software while demonstrators were still present. Encouraged by this result a second trial was conducted, this time without demonstrators present when the students used the courseware. Again the results were encouraging, but the evaluation demonstrated that students had not always succeeded in getting answers to some of their questions. The final section describes a third trial in which a hypertext software adviser was used to provide students with additional help. The evaluation of this trial demonstrates that students found the adviser helpful and were more likely to accept the software and to find answers to their questions.

## Context

Many evaluations of the adoption of courseware have a positive result in that they tend to demonstrate that student learning is either unaffected or improved by the intervention; for

example, see the 'the no significant difference phenomenon' (Russell, 1999; Johnson, Aragon, Shaik and Palma-Rivas, 2000).

However, there have been some more recent ethnographic studies (such as Hara and Kling, 1999) pointing out that students are not always pleased with the learning experience, and may be frustrated by the environment and their inability to talk with someone to solve their problems. In the more extreme cases they see such methods as attempts by universities to absolve themselves of their teaching duties (Noble, 1998). As a result recent research has examined ways in which communication with and between students can be maintained within an online environment (Wegerif, 1998; Arvan, Ory, Bullock, Burnaska and Hanson, 1998). This work adds further evidence to the debate, and introduces a software adviser agent as part of the solution to the problem of maintaining dialogue.

A particular feature of courseware that this work addresses is that of replacing practical laboratories with virtual practicals. There is a body of work in this area which aims at producing virtual or remote practicals (for example, Colwell and Scanlon, 2001) in order to make the experience available to distance students, to share expensive experiments more widely or to make dangerous experiments safer. The focus of this work was concerned with making economies in terms of laboratory equipment, laboratory space and demonstrator time.

### **The SCALPEL courseware**

The general principle in building the SCALPEL (Southampton Computer Assisted Learning Pathology Education Laboratory) courseware was to design an environment to replicate the same six case studies previously delivered in the traditional laboratory. The design constitutes a main window containing these case studies. Hyperlinks were authored into the text to provide access to the same supplementary material previously provided in the lab, that is, pictures, videos and images of microscope work. Additional online background material could easily be referenced and searched when required. A Multiple Choice Question (MCQ) engine was also integrated with the courseware in order to allow the students immediate feedback on whether they were getting the correct answers to the questions in the case study.

The courseware was implemented using Microcosm (Hall, Davis and Hutchings, 1996). The rationale for this decision was mainly due to Microcosm's facilities for integrating with third-party applications; in this experiment two-way integration with the MCQ engine (provided by the STOMP TLTP project) was an important feature, and in the case where a student asked for help from background materials the engine had to be able to follow links back to content. Integration with the Media Viewer was also important. Two other features of Microcosm that were useful were generic links (Fountain, Hall, Heath and Davis, 1990) which allowed the rapid authoring of links from all occurrences of keywords and phrases to appropriate materials, and the computer-link facility which used advanced text search features dynamically to locate suitable materials for the user.

The initial design of SCALPEL was in 1996; at this stage the Web was still in its infancy and browsers had support for little more than rendering of html. The Web was seen as encouraging a didactic view of learning, rather than the student-controlled exploratory style (Crook and Webster, 1997) that was needed for these case studies.

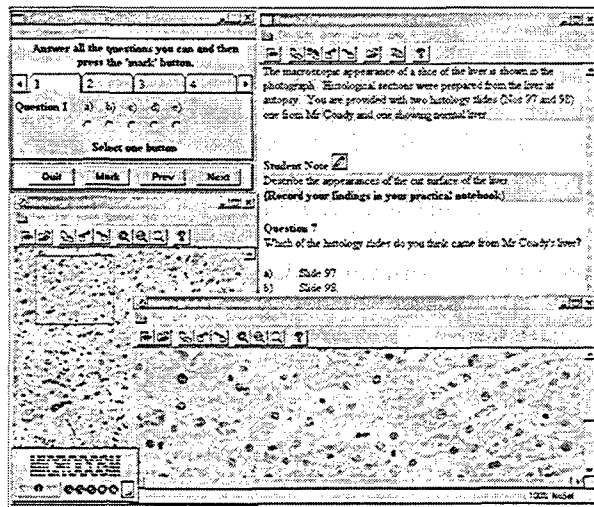


Figure 1: The SCALPEL interface

Figure 1 shows a screenshot of the SCALPEL interface. The case study notes are on the right. The user has followed a link to see two histology slides, and the MCQ engine is ready for the student's answer to the question in the case study.

It would be possible to create a fairly similar learning experience using any of the now widely available virtual learning environments (VLEs), but the choice of environment was fortuitous in making the courseware simple to deliver.

### Evaluation methodology

The evaluations were undertaken in three consecutive years. In each year there were around 160 first-year undergraduate medical students. The unchanged selection procedures and the collection of demographic data gave us confidence that from year to year these students formed a very similar selection of the population, and that it was therefore possible to generalize results from year to year. It is probably reasonable further to generalize the results for the whole body of undergraduate medical students but wider generalization to the population as a whole should only be done with caution.

Quantitative data were collected using questionnaires. These questionnaires asked for answers to factual and non-factual attitudinal information. The latter answers were collected on a five-point Likert scale (Strongly agree, Agree, Neutral, Disagree, Strongly disagree). The sense of the wording of some indicators was reversed to prevent the results being affected by, for example, students agreeing with everything. Complex attitudes were measured using a number of facets of that attitude, and then those facets were averaged. Questionnaires were pre-trialled with colleagues to highlight any problems before they were presented to students. All questionnaires were anonymous.

In the first trial the practical sessions were timetabled, and questionnaires were distributed at the halfway point and at the end of the final practical, so that return rates were 100 per cent ( $N=157$ ). In the second trial students completed the practicals in their own time, and were requested to return questionnaires when they had finished. This led to a much lower response rate ( $N=97/160$ ). In the final trial, in spite of organizing a sweepstake offering two

£50 prizes (students could cut a corner off their questionnaire, write their name on it and add it to the sweepstake, thus maintaining anonymity), the return rate fell even further (N=74/165).

In addition to the quantitative data collection, focus groups were held and students were observed using the system. No attempt was made to quantify the data collected, but rather the results were used to add substance to and to help interpret the results gained from analysis of the questionnaires.

Those who wish to see greater depth of methodology are referred to Michael Kemp's Ph.D. thesis (Kemp, 2000) to see the hypotheses tested, the statistical tests applied and the full experimental results.

### **Online labs: the first evaluation**

The main purpose of the first evaluation was to discover whether there was any measurable change in the effectiveness of learning when lab practicals were replaced with SCALPEL courseware. Clearly an important facet of the introduction of any new teaching method is to convince staff that students would not fare worse as the result of the change. As a sub-agenda we chose to investigate whether there were any groups of students who fared better or worse as the result of the change, where groups were defined by such concepts as demographics, preferred learning style and previous exposure to computers.

For this part of the study twenty-three hypotheses were tested: the fundamental three were about pedagogic value.

1. The educational attainment of pathology on average differs between medical students who work through practicals in the traditional laboratory and those who use SCALPEL.
2. Medical students find SCALPEL an acceptable pedagogic environment.
3. Medical students' acceptance of SCALPEL influences educational attainment.

Hypotheses 4–9 looked at what aspects of students (demographics, learning styles, attitudes, and so on) influenced the pedagogic value of SCALPEL. Hypotheses 10–13 looked at what student characteristics affected their need for demonstrators. Hypotheses 14–18 looked at what characteristics might affect their attitude to computers and hypotheses 19–23 looked at how student characteristics affected learning style.

For the purpose of this study the students were split into two groups. Group A did the first three practicals using the traditional lab and the next three practicals using SCALPEL, whereas Group B did the first three practicals using SCALPEL and the next three in the traditional lab. Demonstrators were present in the lab and in the computer rooms where the SCALPEL practicals were held in the normal timetabled slots.

Educational attainment (ability to recall and apply subject domain knowledge) was measured by examination using MCQ (one mark for right answer, minus one for wrong answer), and short-answer tests. These examinations were taken by all 159 students after the first three practicals and after all practicals were completed. Questionnaires were administered before the start of the first practical, at the end of their final practical using SCALPEL and after all practicals were completed.

The important outcomes of this evaluation were that there was no relationship between educational attainment as measured and the method used for study. Both groups of students scored similar marks at the crossover and at the end, and the marks at the end were an improvement on the marks at the crossover point. Acceptance of SCALPEL was distributed evenly around 'Neutral' and no significant relationship was found between educational attainment and acceptance of SCALPEL.

Other important results were that the students expressed a strong preference for maintaining the presence of demonstrators in the labs when using SCALPEL, and that there was a weak negative relationship between those that expressed a need for demonstrators and expressed acceptance of SCALPEL (that is, those who did not enjoy SCALPEL were more inclined to need demonstrators). These results were further emphasized by the focus groups.

No significant relationships were discovered between students grouped by demographics, learning styles or previous exposure to computers, and their acceptance of SCALPEL or their educational attainment.

Observations indicated that the dynamics of the lab sessions changed with the introduction of SCALPEL. The labs were quiet and the students worked mostly alone, although small ad-hoc groups of two or three students might form to discuss a problem occasionally. Demonstrators were observed moving around the lab to deal with individual queries rather than groups, and as a consequence often answered the same question many times. Few problems were observed with the software although in their first session it was noticed that some students took a while to learn the importance of closing a window after finishing with its contents.

### **Doing away with demonstrators: the second evaluation**

In spite of the strength of feeling expressed by the students in the first trial concerning the need for demonstrator assistance, the team were encouraged by the results showing no change in educational attainment, and decided to continue the experiment by using SCALPEL without demonstrators.

For the second trial, all students were asked to complete the six case studies using SCALPEL. The work was self-paced, meaning they could do it when they liked as no demonstrators would be present in the computer rooms anyway, but tutorials would still be run at the set times and would cover the points that were raised from the last practical that they were expected to have completed.

The important feature of this evaluation was the attempt to discover the methods that students used to solve problems they encountered during practicals. They were asked to note the number of times they would have liked to have consulted a demonstrator, why they wanted to consult a demonstrator, and how they eventually answered their question (whether they used the text book, lecture notes, resources within SCALPEL, asked a student, asked a member of staff, asked in tutorial, or whether they never did answer their question).

Figure 4 below shows the variation of expressed need for demonstrators during SCALPEL sessions, and it is clear that year by year the students showed less need for demonstrators,

in spite of the fact, that in Year 2 at least, nothing had been done to compensate for their absence. This effect is discussed more fully below.

It was clear that the reasons for wishing to consult a demonstrator were nearly all concerned with the pathology content of the practicals and very rarely to do with procedural issues or operation of the software.

Figure 2 shows the methods students used to solve their problems. Text books, tutorial dialogue and peer dialogue are clearly important. Staff were rarely consulted, and the use of the SCALPEL courseware materials is small, indicating that students did not think it likely that answers to their questions would be discovered online.

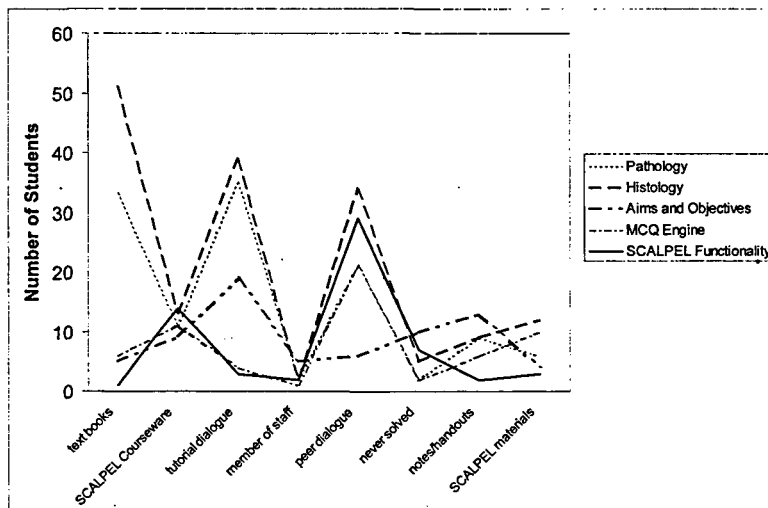


Figure 2: Methods used by students to solve problems in the absence of demonstrators

Students were also asked to suggest how SCALPEL might be improved by ranking the following and by making further suggestions:

- glossary of terms;
- online text books;
- ability to ask questions online and get answers;
- on line Notepad;
- more links to background material;
- optional extra practicals;
- more animations;
- sound clips to introduce sections, such as patient histories.

Students were clear that their preferred improvements to the SCALPEL courseware would involve putting text books online, providing rich linking to the new online materials and the ability to ask questions online. These results were confirmed in focus groups; students

showed little interest in the introduction of multimedia gizmos, in spite of the fact that a few high-quality examples had been created to demonstrate the sort of material that was envisaged.

### **Introducing a hypertext adviser: the third evaluation**

The importance of dialogue with peers and teachers is well documented (see, for example, Laurillard, 2001; Schank and Clearly, 1995; Mayes, 1995). In the earlier versions of SCALPEL the only way that the system entered into any dialogue with the student was via the MCQ engine, which asked the student questions and gave them some feedback on their answers. Clearly this is limited by the fact that the subject of the dialogue is decided by the system and there is no opportunity for the students to ask their own questions. The fact that students were allowed to carry out the SCALPEL practicals at the time of their choice ruled out the use of synchronous communication methods for online dialogue, as there were unlikely to be any staff or many students online at a convenient moment. On the other hand, there has been much work recently on the part that asynchronous communication can play in providing tutorial support and dialogue (such as Masterton, 1998; Stratfold, 1998). Many of these approaches are based on work on 'Answer Gardens' (Ackerman and Malone, 1990). In this approach students pose questions, and the answers to these questions are collected in a kind of extended FAQ, organized by topic so that students can locate areas of similar questions to their own. As more questions are posed and answered, the answer garden grows.

The final experiment set out to discover whether the introduction of a richer environment supported by an online 'demonstrator agent' would improve attitudes to SCALPEL. In this experiment all 165 students completed the same six practicals, but an experimental group of 40 volunteers were given additional access to an enhanced MCQ engine which provided access to the demonstrator agent. Of these, 28 students from the experimental group and 46 from the control group returned questionnaires.

Figure 3 shows a typical screen shot of the demonstrator agent in use. The student wishes to answer question 2, but has a question. The demonstrator agent is aware of the context (the question the student is trying to answer) and is therefore able to list all other questions that have previously been asked in this context (as well as offering the student the opportunity to consult other contexts). In the example the student has selected to see the answer to one of these questions. Alternatively the student could have submitted a new query. This would have been answered in due course by a domain expert, and thus added to the list of previously asked questions. In this way the answer garden grows.

In the example the student has read the answer, but is perhaps not clear of the exact meaning of one of the terms that is used in the answer, so has pressed the Show Links button. This has sent a query to Microcosm asking it to show any links it has in its database which are anchored on terms which appear in the answer to the question. The screenshot shows Microcosm returning with two links to background materials (in this case a glossary of terms, but it might have been an online text book) that might be useful.

The results of the evaluation were encouraging. Figure 4 shows the decreasing need that students as a whole show for consulting demonstrators. This was directly linked to a year by year improvement in the acceptance of SCALPEL.



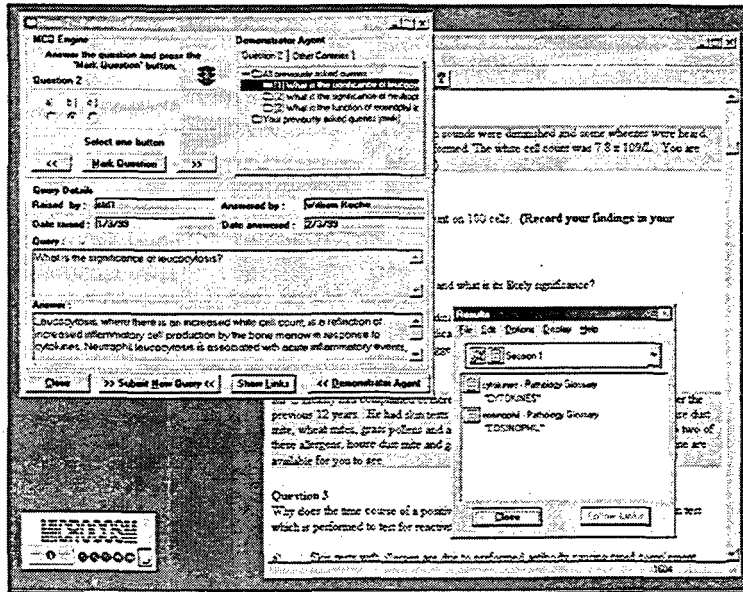


Figure 3: Screenshot of SCALPEL's demonstrator agent in use

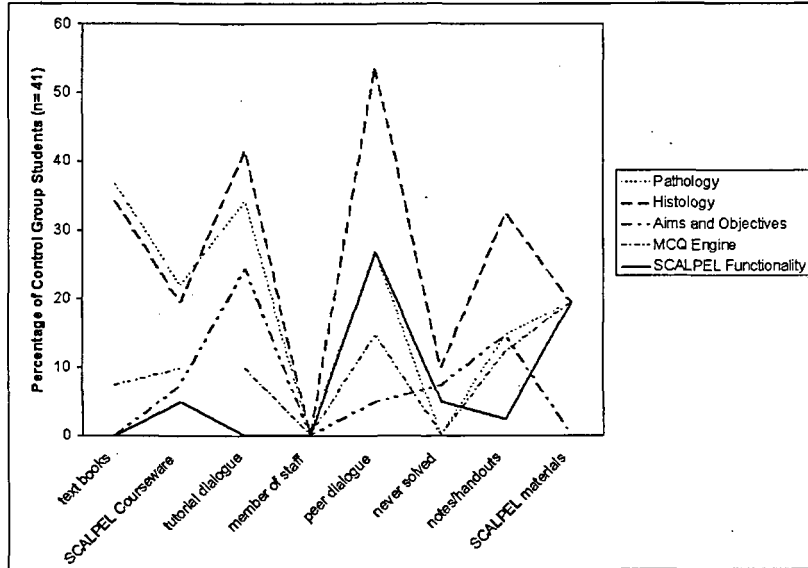


Figure 4: Students' expressed need for demonstrators during SCALPEL sessions

It is not clear why this year on year improvement occurred. One possible interpretation is that in the late 1990s computers were becoming increasingly widespread (this was confirmed in their responses to questions about previous exposure) and in particular the Web was becoming a standard source of reference. A generation of students brought up with expectations of learning from and referring to online materials is more likely to find courseware acceptable.

There was little difference between the experimental group and the control group in the reasons expressed for wishing to consult a demonstrator. However, the control group expressed an average need of seven demonstrator consults per student compared with an average of five per student for the experiment group.

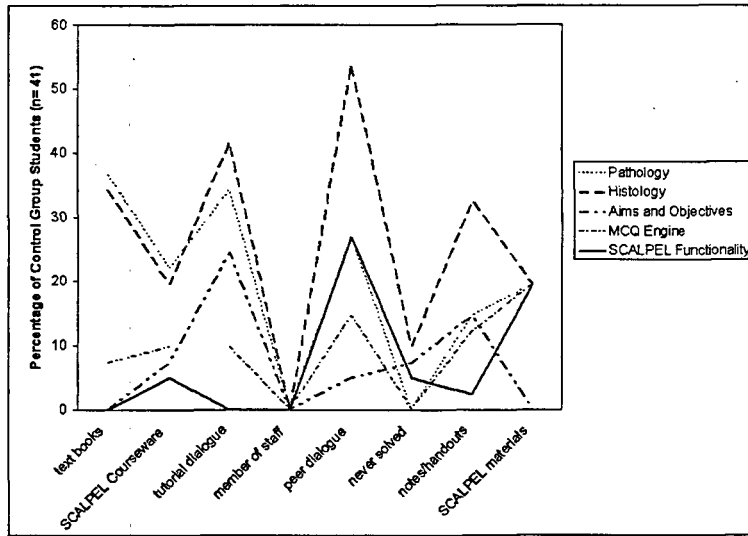


Figure 5: How the control group solved problems

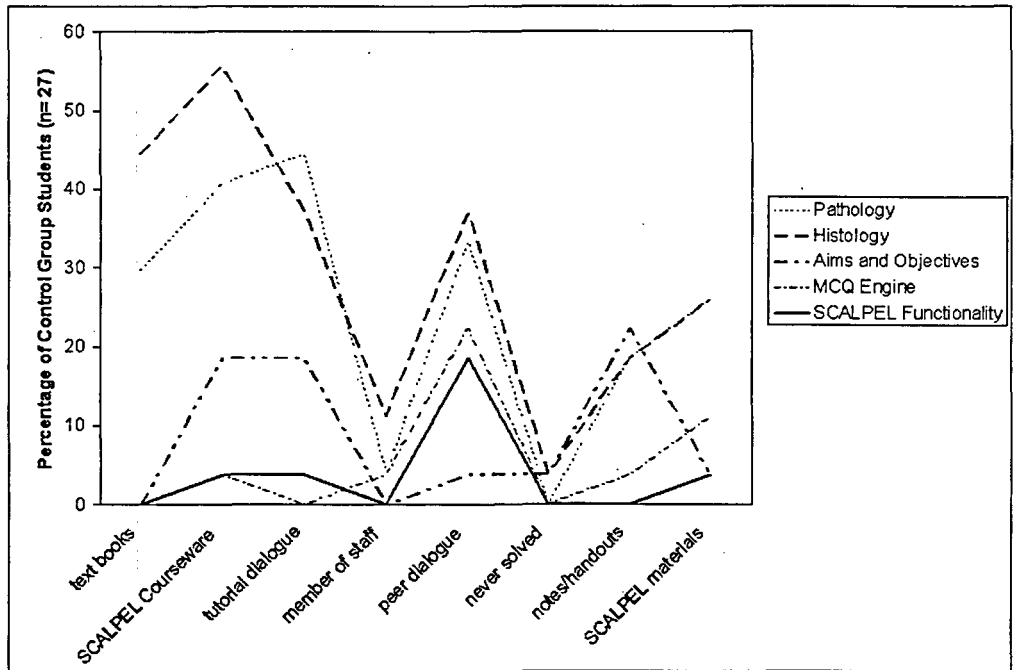


Figure 6: How the group with access to the demonstrator agent solved problems

The pattern for solving problems changed considerably as shown in Figures 5 and 6. In particular we see that the SCALPEL courseware becomes much more used, questions to members of staff become used a little (by submitting questions via the agent) and the number of questions never solved is considerably reduced.

There was a significant improvement in attitude to the use of SCALPEL recorded by those who had access to the demonstrator agent. This is attributed to a reduction in the average number of required demonstrator consults relative to the control group, and an increase in problems solved as a consequence of the demonstrator agent facilitating the resolution of misconceptions with the pathology material.

## Conclusions

This series of evaluation experiments has demonstrated that there was no significant change in student learning as the result of replacing traditional labs with virtual labs, even with the earliest group who were most averse to using courseware.

An unexpected result of this work was the fact that, over the three years, the first-year students were increasingly accepting of the use of courseware and found it less important to consult demonstrators. This was independent of any changes we made in the courseware. It is possible that this effect could be in some way linked to the reducing percentage of students that answered the surveys, but is more probably a reflection of the way society is changing. Over the period of this study in the late 1990s PCs, and in particular the Web and email, became standard tools for most educated citizens of the UK, and as a consequence the students were more accustomed to reading, learning from and interacting with computers.

In spite of the general improvement in acceptance of courseware, there was a general preference for access to demonstrators. This is not surprising; models of learning based on constructivist views (Brown, Metz and Campione, 1996) and conversational frameworks (Laurillard, 2001) would lead one to expect that removing demonstrator assistance would lead to poorer learning, as reported by our students. The encouraging result was that the acceptance of the use of courseware was significantly improved by the reintroduction of a relatively simple software agent, which provided limited dialogue via an answer garden and dynamic access to suitable in-context links.

It is testament to the success of this work that all pathology practicals at the University of Southampton were subsequently delivered to students through the SCALPEL environment.

## References

- Ackerman, M. S. and Malone, T. W. (1990), 'Answer Garden: a tool for growing organizational memory', in the *Proceedings of the ACM Conference on Office Information Systems*, New York: ACM, 31-9.
- Arvan, L., Ory, J. C., Bullock, C. D., Burnaska, K. K. and Hanson, M. (1998), 'The SCALE efficiency projects', *JALN*, 2 (2), [http://www.aln.org/alnweb/journal/jaln\\_vol2issue2.htm](http://www.aln.org/alnweb/journal/jaln_vol2issue2.htm).
- Brown, A. L., Metz, K. E. and Campione, J. C. (1996), 'Social interaction and individual

understanding in a community of learners: the influence of Piaget and Vygotsky', in A. Tryphon, and J. Vonèche (eds.), *Piaget-Vygotsky: The Social Genesis of Thought*, Brighton and New York: Psychology Press, 145–70.

Colwell, C and Scanlon, E. (2001), 'The PEARL project: remote, collaborative experimentation for science students', in *The Proceedings of CAL 2001*, Warwick: Warwick University.

Crook, C. and Webster, D. S. (1997), 'Designing for informal undergraduate computer mediated communication', *Active Learning*, 7.

Davis, H. C., Hutchings, G. A. and Hall, W. (1993), 'A framework for delivering large-scale hypermedia learning material', in H. Maurer (ed.), *Educational Multimedia and Hypermedia Annual 1993*, Proceedings of ED-MEDIA '93, Orlando, FL: AACE, 115–22.

Fountain, A. M., Hall, W., Heath, I. and Davis, H. C. (1990), 'MICROCOSM: an open model for hypermedia with dynamic linking', in A. Rizk, N. Streitz and J. Andre (eds.), *Hypertext: Concepts, Systems and Applications*, The Proceedings of the European Conference on Hypertext, INRIA, France, November 1990, Cambridge: Cambridge University Press.

Hall, W., Davis, H. C. and Hutchings, G. A. (1996), *Rethinking Hypermedia: The Microcosm Approach*, Vancouver, BC: Kluwer Academic Publishers.

Hara, N. and Kling, R. (1999), 'Students' frustrations with a Web-based distance education course', *First Monday*, 4 (12),  
[http://www.firstmonday.dk/issuues/issue4\\_12/haralindex.html](http://www.firstmonday.dk/issuues/issue4_12/haralindex.html)

Johnson, S. D., Aragon, S. R., Shaik, N. and Palma-Rivas, N. (2000), 'Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments', *Journal of Interactive Learning Research*, 11 (1), 29–49.

Kemp, M. W. (2000), 'From classroom tutor to hypermedia advisor: a case study in medical education', Ph.D. thesis, University of Southampton.

Laurillard, D. (2001), *Rethinking University Teaching: A Framework for the Effective Use of Educational Technology*, 2nd edn, London: Routledge.

Masterton, S. (1998), 'The virtual participant: a tutor's assistant for electronic conferencing', in M. Eisenstadt and T. Vincent (eds.), *The Knowledge Web: Learning and Collaborating on the Net*, London: Kogan Page, 249–65.

Mayes, J. T. (1993), 'Commentary: impact of cognitive theory on the practice of courseware authoring', *Journal of Computer Assisted Learning*, 9, 222–8.

Mayes, J. T. (1995), 'Learning technology and Groundhog Day', in W. Strang and D. Slater (eds.), *Hypermedia at Work: Practice and Theory in Higher Education*, Canterbury: University of Kent Press.

McCullagh, P. and Roche, W. R. (1992), 'First year pathology practicals: the Southampton Experience', *Journal of Pathology*, 169 (supplement 148A).

Noble, D. F. (1998), 'Digital diploma mills: the automation of higher education', *First Monday*, 3 (1), [http://firstmonday.org/issuues/issue3\\_1/noble/index.html](http://firstmonday.org/issuues/issue3_1/noble/index.html)

Russell, T. L. (1999), *The No Significant Difference Phenomenon*, New Brunswick: IDECC, and also see the accompanying Website: <http://teleeducation.nb.ca/nosignificantdifference/>

Schank, R. C and Clearly, C. (1995), *Engines for Education*, Hillsdale, NJ: Lawrence Erlbaum Associates.

Stratfold, M. (1998), 'Promoting learner dialogues on the Web', in M. Eisenstadt and T. Vincent (eds.), *The Knowledge Web: Learning and Collaborating on the Net*, London: Kogan Page, 119–34.

Wegerif, R. (1998), 'The social dimension of asynchronous learning networks', *Journal of Asynchronous Learning Networks*, 2 (1), [http://www.aln.org/alnweb/journal/vol2\\_issue1/wegerif.htm](http://www.aln.org/alnweb/journal/vol2_issue1/wegerif.htm)